JC09 Rec'd PCT/PTO 28 FEB 2001
09/786009

Practitioner's Docket No. NEB-150PUS

CHAPTER II

Preliminary Classification:

Proposed Class:

Subclass:

NOTE: "All applicants are requested to include a preliminary classification on newly filed patent applications. The preliminary classification, preferably class and subclass designations, should be identified in the upper right-hand corner of the letter of transmittal accompanying the application papers, for example 'Proposed Class 2, subclass 129.'" M.P.E.P., § 601, 7th ed.

TRANSMITTAL LETTER
TO THE UNITED STATES ELECTED OFFICE (EO/US)

(ENTRY INTO U.S. NATIONAL PHASE UNDER CHAPTER II)

PCT/US99/22776	30 September 1999	30 September 1998
INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED

Intein Mediated Peptide Ligation

TITLE OF INVENTION

Ming-Qun XU, Thomas C. EVANS

APPLICANT(S)

Box PCT
Assistant Commissioner for Patents
Washington D.C. 20231

ATTENTION: EO/US

CERTIFICATION UNDER 37 C.F.R. § 1.10*

*(Express Mail label number is **mandatory**.)*
(Express Mail certification is optional.)

I hereby certify that this Transmittal Letter and the papers indicated as being transmitted therewith is being deposited with the United States Postal Service on this date 28 February 2001 in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number EL010489827US, addressed to the: Assistant Commissioner for Patents, Washington, D.C. 20231.

Melissa A. Jackson
(type or print name of person mailing paper)

Signature of person mailing paper

***WARNING:** Certificate of mailing (first class) or facsimile transmission procedures of 37 C.F.R. § 1.8 cannot be used to obtain a date of mailing or transmission for this correspondence.*

****WARNING:** Each paper or fee filed by "Express Mail" **must** have the number of the "Express Mail" mailing label placed thereon prior to mailing. 37 C.F.R. § 1.10(b).*
*"Since the filing of correspondence under § 1.10 without the Express Mail mailing label thereon is an oversight that can be avoided by the exercise of reasonable care, requests for waiver of this requirement will **not** be granted on petition." Notice of Oct. 24, 1996, 60 Fed. Reg. 56,439, at 56,442.*

NOTE: To avoid abandonment of the application, the applicant shall furnish to the USPTO, not later than 20 months from the priority date: (1) a copy of the international application, unless it has been previously communicated by the International Bureau or unless it was originally filed in the USPTO; and (2) the basic national fee (see 37 C.F.R. § 1.492(a)). The 30-month time limit may not be extended. 37 C.F.R. § 1.495.

WARNING: *Where the items are those which can be submitted to complete the entry of the international application into the national phase are subsequent to 30 months from the priority date the application is still considered to be in the international state and if mailing procedures are utilized to obtain a date the express mail procedure of 37 C.F.R. § 1.10 must be used (since international application papers are not covered by an ordinary certificate of mailing—See 37 C.F.R. § 1.8.*

NOTE: Documents and fees must be clearly identified as a submission to enter the national state under 35 U.S.C. § 371 otherwise the submission will be considered as being made under 35 U.S.C. § 111. 37 C.F.R. § 1.494(f).

I. Applicant herewith submits to the United States Elected Office (EO/US) the following items under 35 U.S.C. § 371:

a. ☒ This express request to immediately begin national examination procedures (35 U.S.C. § 371(f)).

b. ☒ The U.S. National Fee (35 U.S.C. § 371(c)(1)) and other fees (37 C.F.R. § 1.492) as indicated below:

2. Fees

CLAIMS FEE	(1) FOR	(2) NUMBER FILED	(3) NUMBER EXTRA	(4) RATE	(5) CALCULATIONS
☐*	TOTAL CLAIMS	15 −20=	0.00	× $18.00=	$ 0.00
	INDEPENDENT CLAIMS	3 −3=	0.00	× $80.00=	0.00
	MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ $270.00	270.00
BASIC FEE**	☒ U.S. PTO WAS INTERNATIONAL PRELIMINARY EXAMINATION AUTHORITY Where an International preliminary examination fee as set forth in § 1.482 has been paid on the international application to the U.S. PTO: ☐ and the international preliminary examination report states that the criteria of novelty, inventive step (non-obviousness) and industrial activity, as defined in PCT Article 33(1) to (4) have been satisfied for all the claims presented in the application entering the national stage (37 C.F.R. § 1.492(a)(4)) .... $100.00 ☒ and the above requirements are not met (37 C.F.R. § 1.492(a)(1)) .... $690.00 ☐ U.S. PTO WAS NOT INTERNATIONAL PRELIMINARY EXAMINATION AUTHORITY Where no international preliminary examination fee as set forth in § 1.482 has been paid to the U.S. PTO, and payment of an international search fee as set forth in § 1.445(a)(2) to the U.S. PTO: ☐ has been paid (37 C.F.R. § 1.492(a)(2)) .... $710.00 ☐ has not been paid (37 C.F.R. § 1.492(a)(3)) ..... $1000.00 ☐ where a search report on the international application has been prepared by the European Patent Office or the Japanese Patent Office (37 C.F.R. § 1.492(a)(5)) $860.00				710.00
	Total of above Calculations				= 980.00
SMALL ENTITY	Reduction by 1/2 for filing by small entity, if applicable. Affidavit must be filed also. (note 37 C.F.R. § 1.9, 1.27, 1.28)				− 490.00
	Subtotal				490.00
	Total National Fee				$ 490.00
	Fee for recording the enclosed assignment document $40.00 (37 C.F.R. § 1.21(h)). (See Item 13 below). See attached "ASSIGNMENT COVER SHEET".				40.00
TOTAL	Total Fees enclosed				$ 530.00

*See attached Preliminary Amendment Reducing the Number of Claims.

☒ Attached is a ☒ check ☐ money order in the amount of $ 530.00

☐ Authorization is hereby made to charge the amount of $ ________

☒ to Deposit Account No. 14-0740

☐ to Credit card as shown on the attached credit card information authorization form PTO-2038.

WARNING: *Credit card information should* ***not*** *be included on this form as it may become public.*

☒ Charge any additional fees required by this paper or credit any overpayment in the manner authorized above.

A duplicate of this paper is attached.

****WARNING:*** *"To avoid abandonment of the application the applicant shall furnish to the United States Patent and Trademark Office not later than the expiration of 30 months from the priority date: * * * (2) the basic national fee (see § 1.492(a)). The 30-month time limit may not be extended." 37 C.F.R. § 1.495(b).*

WARNING: *If the translation of the international application and/or the oath or declaration have not been submitted by the applicant within thirty (30) months from the priority date, such requirements may be met within a time period set by the Office. 37 C.F.R. § 1.495(b)(2). The payment of the surcharge set forth in § 1.492(e) is required as a condition for accepting the oath or declaration later than thirty (30) months after the priority date. The payment of the processing fee set forth in § 1.492(f) is required for acceptance of an English translation later than thirty (30) months after the priority date. Failure to comply with these requirements will result in abandonment of the application. The provisions of § 1.136 apply to the period which is set. Notice of Jan. 3, 1993, 1147 O.G. 29 to 40.*

3. ☒ A copy of the International application as filed (35 U.S.C. § 371(c)(2)):

NOTE: Section 1.495 (b) was amended to require that the basic national fee and a copy of the international application must be filed with the Office by 30 months from the priority date to avoid abandonment. "The International Bureau normally provides the copy of the international application to the Office in accordance with PCT Article 20. At the same time, the International Bureau notifies applicant of the communication to the Office. In accordance with PCT Rule 47.1, that notice shall be accepted by all designated offices as conclusive evidence that the communication has duly taken place. Thus, if the applicant desires to enter the national stage, the applicant normally need only check to be sure the notice from the International Bureau has been received and then pay the basic national fee by 30 months from the priority date." Notice of Jan. 7, 1993, 1147 O.G. 29 to 40, at 35-36. See item 14c below.

a. ☐ is transmitted herewith.

b. ☒ is not required, as the application was filed with the United States Receiving Office.

c. ☐ has been transmitted

i. ☐ by the International Bureau.
Date of mailing of the application (from form PCT/1B/308):

________.

ii. ☐ by applicant on ________. *(Date)*

4. ☒ A translation of the International application into the English language (35 U.S.C. § 371(c)(2)):

a. ☐ is transmitted herewith.

b. ☒ is not required as the application was filed in English.

c. ☐ was previously transmitted by applicant on ________. *(Date)*

d. ☐ will follow.

5. ☒ Amendments to the claims of the International application under PCT Article 19 (35 U.S.C. § 371(c)(3)):

NOTE: *The Notice of January 7, 1993 points out that 37 C.F.R. § 1.495(a) was amended to clarify the existing and continuing practice that PCT Article 19 amendments must be submitted by 30 months from the priority date and this deadline may not be extended. The Notice further advises that: "The failure to do so will not result in loss of the subject matter of the PCT Article 19 amendments. Applicant may submit that subject matter in a preliminary amendment filed under section 1.121. In many cases, filing an amendment under section 1.121 is preferable since grammatical or idiomatic errors may be corrected." 1147 O.G. 29-40, at 36.*

a. ☐ are transmitted herewith.

b. ☐ have been transmitted

i. ☐ by the International Bureau.
Date of mailing of the amendment (from form PCT/1B/308):
________.

ii. ☐ by applicant on ________. *(Date)*

c. ☒ have not been transmitted as

i. ☒ applicant chose not to make amendments under PCT Article 19.
Date of mailing of Search Report (from form PCT/ISA/210.):
10 April 2000.

ii. ☐ the time limit for the submission of amendments has not yet expired. The amendments or a statement that amendments have not been made will be transmitted before the expiration of the time limit under PCT Rule 46.1.

6. ☒ A translation of the amendments to the claims under PCT Article 19 (38 U.S.C. § 371(c)(3)):

a. ☐ is transmitted herewith.

b. ☐ is not required as the amendments were made in the English language.

c. ☒ has not been transmitted for reasons indicated at point 5(c) above.

7. ☒ A copy of the international examination report (PCT/IPEA/409)

☐ is transmitted herewith.

☒ is not required as the application was filed with the United States Receiving Office.

8. ☒ Annex(es) to the international preliminary examination report

a. ☐ is/are transmitted herewith.

b. ☒ is/are not required as the application was filed with the United States Receiving Office.

9. ☒ A translation of the annexes to the international preliminary examination report

a. ☐ is transmitted herewith.

b. ☒ is not required as the annexes are in the English language.

10. ☒ An oath or declaration of the inventor (35 U.S.C. § 371(c)(4)) complying with 35 U.S.C. § 115

a. ☐ was previously submitted by applicant on ________.
Date

b. ☒ is submitted herewith, and such oath or declaration

i. ☒ is attached to the application.

ii. ☐ identifies the application and any amendments under PCT Article 19 that were transmitted as stated in points 3(b) or 3(c) and 5(b); and states that they were reviewed by the inventor as required by 37 C.F.R. § 1.70.

c. ☐ will follow.

II. Other document(s) or information included:

11. ☐ An International Search Report (PCT/ISA/210) or Declaration under PCT Article 17(2)(a):

a. ☐ is transmitted herewith.

b. ☐ has been transmitted by the International Bureau.
Date of mailing (from form PCT/IB/308): ________.

c. ☐ is not required, as the application was searched by the United States International Searching Authority.

d. ☐ will be transmitted promptly upon request.

e. ☐ has been submitted by applicant on ________.
Date

12. ☒ An Information Disclosure Statement under 37 C.F.R. §§ 1.97 and 1.98:

a. ☐ is transmitted herewith.

Also transmitted herewith is/are:

☐ Form PTO-1449 (PTO/SB/08A and 08B).

☐ Copies of citations listed.

b. ☒ will be transmitted within THREE MONTHS of the date of submission of requirements under 35 U.S.C. § 371(c).

c. ☐ was previously submitted by applicant on ________.
Date

13. ☒ An assignment document is transmitted herewith for recording.

A separate ☐ "COVER SHEET FOR ASSIGNMENT (DOCUMENT) ACCOMPANYING NEW PATENT APPLICATION" or ☒ FORM PTO 1595 is also attached.

New England Biolabs, Inc.

14. ☐ Additional documents:

a. ☐ Copy of request (PCT/RO/101)

b. ☐ International Publication No. __________

i. ☐ Specification, claims and drawing

ii. ☐ Front page only

c. ☐ Preliminary amendment (37 C.F.R. § 1.121)

d. ☐ Other

15. ☐ The above checked items are being transmitted

a. ☐ before 30 months from any claimed priority date.

b. ☐ after 30 months.

16. ☐ Certain requirements under 35 U.S.C. § 371 were previously submitted by the applicant on ____________, namely:

AUTHORIZATION TO CHARGE ADDITIONAL FEES

WARNING: *Accurately count claims, especially multiple dependant claims, to avoid unexpected high charges if extra claims are authorized.*

NOTE: *"A written request may be submitted in an application that is an authorization to treat any concurrent or future reply, requiring a petition for an extension of time under this paragraph for its timely submission, as incorporating a petition for extension of time for the appropriate length of time. An authorization to charge all required fees, fees under § 1.17, or all required extension of time fees will be treated as a constructive petition for an extension of time in any concurrent or future reply requiring a petition for an extension of time under this paragraph for its timely submission. Submission of the fee set forth in § 1.17(a) will also be treated as a constructive petition for an extension of time in any concurrent reply requiring a petition for an extension of time under this paragraph for its timely submission." 37 C.F.R. § 1.136(a)(3).*

NOTE: *"Amounts of twenty-five dollars or less will not be returned unless specifically requested within a reasonable time, nor will the payer be notified of such amounts; amounts over twenty-five dollars may be returned by check or, if requested, by credit to a deposit account." 37 C.F.R. § 1.26(a).*

☒ Please charge, in the manner authorized above, the following additional fees that may be required by this paper and during the entire pendency of this application:

☒ 37 C.F.R. § 1.492(a)(1), (2), (3), and (4) (filing fees)

WARNING: *Because failure to pay the national fee within 30 months without extension (37 C.F.R. § 1.495(b)(2)) results in abandonment of the application, it would be best to always check the above box.*

09/786009
JC02 Rec'd PCT/PTO 2 8 FEB 2001

☒ 37 C.F.R. § 1.492(b), (c) and (d) (presentation of extra claims)

NOTE: Because additional fees for excess or multiple dependent claims not paid on filing or on later presentation must only be paid or these claims cancelled by amendment prior to the expiration of the time period set for response by the PTO in any notice of fee deficiency (37 C.F.R. § 1.492(d)), it might be best not to authorize the PTO to charge additional claim fees, except possible when dealing with amendments after final action.

☐ 37 C.F.R. § 1.17 (application processing fees)

☐ 37 C.F.R. § 1.17(a)(1)-(5) (extension fees pursuant to § 1.136(a).

☐ 37 C.F.R. § 1.18 (issue fee at or before mailing of Notice of Allowance, pursuant to 37 C.F.R. § 1.311(b))

NOTE: Where an authorization to charge the issue fee to a deposit account has been filed before the mailing of a Notice of Allowance, the issue fee will be automatically charged to the deposit account at the time of mailing the notice of allowance. 37 C.F.R. § 1.311(b).

NOTE: 37 C.F.R. § 1.28(b) requires "Notification of any change in loss of entitlement to small entity status must be filed in the application . . . prior to paying, or at the time of paying . . . issue fee." From the wording of 37 C.F.R. § 1.28(b): (a) notification of change of status must be made even if the fee is paid as "other than a small entity" and (b) no notification is required if the change is to another small entity.

☐ 37 C.F.R. § 1.492(e) and (f) (surcharge fees for filing the declaration and/or filing an English translation of an International Application later than 30 months after the priority date).

SIGNATURE OF PRACTITIONER

Reg. No.: 30901

Gregory D. Williams
General Counsel

Tel. No.: (978) 927-5054 X:292

(*type or print name of practitioner*)

New England Biolabs, Inc.
32 Tozer Road

Customer No.:

P.O. Address
Beverly, MA 01915

INTEIN MEDIATED PEPTIDE LIGATION

BACKGROUND OF THE INVENTION

Genetic engineering is a powerful approach to the manipulation of proteins. However, genetic methodologies are constrained by the use of only naturally coded amino acids. Furthermore, cytotoxic proteins are difficult to obtain by expression and isolation from a living source, since the expression of the toxic protein can result in death of the host.

To some extent, protocols have been developed to circumvent these problems, for example, total chemical synthesis (Kent, S. B. (1988) *Ann. Rev. Biochem.* 57:957-989), use of misacylated tRNAs (Noren, et al., (1989) *Science* 244:182-188), and semi-synthetic techniques (reviewed in Offord, R. (1987) *Protein Eng.* 1:151-157; Roy. et al. (1994) *Methods in Enzymol.* 231:194-215; Wallace, C. J. (1993) *FASEB* 7:505-515). However, all of these procedures are limited by either the size of the fragment which can be generated or by low reaction yield.

It would therefore be desirable to develop a high-yield, semi-synthetic technique to allow *in vitro* fusion of a synthetic protein or peptide fragment to an expressed protein without limitation as to the size of the fused fragments.

Likewise, in order to produce cytotoxic proteins, it would be desirable to develop a method of fusing a synthetic fragment, *in vitro*, to an inactive, expressed protein, so as to restore protein activity post-production from the host.

The modified Sce VMA intein has been used to generate thioester-tagged proteins for use in ligation (Example 19, U.S.S.N. 08/811,492, filed June 16, 1997; Chong, (1996) *J. Biol. Chem.*, 271(36):22159-22168; Chong, (1997) *Gene*, 192:271-281; and Muir, et al. (1998) *Proc. Natl. Acad. Sci* USA 95:6705-6710).

Some disadvantages have been low yields due to poor cleavage of the Sce VMA intein with thiol-reagents that are optimum for ligation, the need for large peptide quantities due to on-column reactions, the use of odoriferous reagents, and/or low protein yields due to the use of a large, eukaryotic intein.

SUMMARY OF THE INVENTION

In accordance with the present invention, there is provided a method for producing a semi-synthetic fusion protein *in vitro*, comprising the steps of producing a target protein fused to a protein splicing element (an intein) and selectively cleaving the fusion and ligating a synthetic

protein or peptide at the C-terminal thioester of the target protein, which overcome many of the disadvantages and problems noted above. Specifically, the present invention has higher yields due to better thiol-induced cleavage with thiol reagents which have been optimized for the ligation reaction. Off-column ligation allows for sample concentration as well as the use of less peptide. In a particularly preferred embodiment, thiol reagents such as 2-mercaptoethanesulfonic acid (MESNA), which is an odorless thiol-reagent, is used for cleavage and ligation along with the Mxe intein, which is from a bacterial source and often expresses better in bacterial cells. Furthermore, the present invention allows peptides to be directly ligated to the thioester bond formed between an intein and the target protein. The present invention also provides a method for producing a cytotoxic protein, comprising the steps of producing a truncated, inactive form of the protein *in vivo* which is fused to a protein splicing element, and selectively cleaving the fusion and ligating a synthetic protein or peptide at a C-terminal thioester of the target protein to restore the activity of the native cytotoxic protein. Recombinant vectors for producing such cleavable fusion proteins are also provided.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a flow diagram depicting the chemical reactions which enable intein-mediated peptide ligation. The

thioester generated at the C-terminus of the target protein during IMPACT™ purification was used in a 'native chemical ligation' reaction. This allowed the ligation of a synthetic peptide to a bacterially expressed protein. A typical ligation reaction involved the expression of the target protein-intein-CBD fusion followed by binding to a chitin resin. A thiol reagent induced cleavage of the intein. The target was eluted from the chitin resin and a synthetic peptide was added. The ligation reaction proceeded overnight.

Figure 2 is a gel depicting the results of cleavage and ligation reactions using various thiols. Cleavage and ligation reactions with different thiols visualized on 10-20% Tricine gels. MYB (a fusion protein of maltose binding protein-Sce VMA intein (N454A)-chitin binding domain) and MXB (a fusion protein of maltose binding protein-Mxe GyrA (N198A) intein-chitin binding domain) were incubated overnight at 4°C with various thiols (50 mM) in 150 mM Tris, 100 mM NaCl, pH 8 in the presence of a 30 amino acid peptide with an N-terminal cysteine. The peptide ligates to the C-terminus of MBP. Lanes 1-5 ligation with MYB. Lane 1 no thiol. Lane 2 dithiothreitol. Lane 3 2-mercaptoethanesulfonic acid. Lane 4 3-mercaptopropionic acid. Lane 5 thiophenol. Lanes 6-10 ligation with MXB. Lane 6 no thiol. Lane 7 dithiothreitol. Lane 8 2-mercaptoethanesulfonic acid. Lane 9 3-mercaptopropionic acid. Lane 10 thiophenol.

Figure 3 is a gel depicting direct ligation of a peptide to the thioester formed between the Sce VMA intein and maltose binding protein. SDS-PAGE of direct ligation reaction with a 10-20% Tricine gel. Lane 1: a precursor protein (MYBleu) consisting of maltose binding protein-Sce VMA1 intein-chitin binding domain was heated to >95°C for 5 minutes in a buffer of 50 mM Trizma base, pH 8.5 containing 100 mM NaCl, 1% SDS, and mM tris-(2-carboxyethyl)phosphine (TCEP) followed by overnight incubation at room temperature. The precursor (MYBleu) is visible along with the Sce VMA1 intein (Y) and maltose binding protein (M), which are cleavage products. Lane 2: the precursor protein was subjected to the same conditions as described in Lane 1 except that the 30 amino acid peptide (1 mM) was added. The precursor (MYB) and cleavage products (Y and M) are visible along with the ligation product (M+30mer) formed when the 30 amino acid peptide fuses to maltose binding protein.

Figure 4 is a diagram depicting the pTXB1 expression vector of Example I (SEQ ID NO:7 and SEQ ID NO:8).

Figure 5 is the DNA sequence of pTXB1 (SEQ ID NO:5).

Figure 6 is a gel depicting the results of the *HpaI* protein ligation reaction. Protein ligation reactions examined on 10-20% Tricine gels. Lane 1: clarified cells extract after IPTG (0.5 mM) induction of ER2566 cells containing the pTXB2-*HpaI*

plasmid. The fusion protein of *HpaI*$_{223}$-Mxe GyrA-intein-CBD (52 kDa) is visible. Lane 2: cell extract as in Lane 1 after passage over a chitin column, which results in the binding of the fusion protein. Lane 3: *HpaI*$_{223}$ (25.7 kDa) after cleavage from the fusion protein by addition of MESNA. Lane 4: ligation product of *HpaI*$_{223}$ (0.2 mg/mL) with 1 mM of a 31 amino acid peptide (ligation product 29.6 kDa), representing the residues necessary to generate full length *HpaI*, after overnight incubation at 4°C. Lane 5: full length *HpaI* from a recombinant source (29.6 kDa) containing BSA (66 kDa) and two impurities.

Figure 7 is a western blot of various proteins ligated to a biotinylated peptide. Proteins purified with the Mxe GyrA IMPACT™ derivative were ligated to a synthetic peptide which contained an antibody recognition sequence.

DETAILED DESCRIPTION OF THE INVENTION

The ligation methods of the present invention are based on the discovery that a cysteine or peptide fragment containing an N-terminal cysteine may be fused, *in vitro*, to a bacterially expressed protein produced by thiol-induced cleavage of an intein (U.S. Patent No. 5,496,714; Example 19 of U.S.S.N. 08/811,492 filed June 16, 1997; Chong, et al., (1996) *supra* and Chong, et al., (1997) *supra*.

The ligation procedure disclosed herein utilizes a protein splicing element, an intein (Perler, et al., (1994) *Nucleic Acids Res.* 22:1125-1127) to precisely create a thioester at the C-terminal α-carbon of an expressed protein. This reactive thioester could be present between the target protein and intein or generated by the addition of a thiol reagent. Previously the generation such a thioester was described using an intein (CIVPS) that was modified to undergo thiol inducible cleavage at its N-terminal junction in the presence of thiol reagent dithiothreitol (DTT) (Chong, et al. (1997) *supra*; Comb, et.al. U.S. Patent No. 5,496,714). This C-terminal thioester was previously used in a 'native chemical ligation' type reaction (Dawson, et al., (1994) *Science* 266:776-779) to fuse ^{35}S-cysteine or a peptide fragment containing an N-terminal cysteine to a bacterially expressed protein (Example 19, Comb, et.al. U.S. Patent No. 5,834,247, Chong (1996) *supra* and Chong (1997) *supra*.

The ligation method of the instant invention begins with the purification of the thioester-tagged target protein using an intein as described (Chong, et.al. (1997) *supra*). The direct ligation method of the instant invention begins with the isolation of a precursor composed of the target protein-intein-CBD. In one preferred embodiment, the host cell is bacterial. In other embodiments the host cell may be yeast, insect, or mammalian. A cysteine thiol at the N-terminus of a synthetic peptide nucleophilicly attacks a thioester present

on the freshly isolated C-terminal α-carbon of the target protein or directly attacks the thioester present between the target protein and intein. This initially generates a thioester between the two reactants which spontaneously rearranges into a native peptide bond (Figure 1).

In order to optimize the ligation efficiency so that greater than 90% of the bacterially expressed target protein can be fused to the synthetic peptide or protein, specific thiol reagents and inteins are screened. In a preferred embodiment, the intein may be any CIVPS, such as *Sce* VMA, *Mxe* GyrA or derivatives of mutants thereof, and the thiol reagent is 2-mercapto-ethanesulfonic acid, thiophenol, DTT, or 3-mercaptopropionic acid (Comb, et al., U.S. Patent No. 5,496,714; U.S. Patent No. 5,834,247).

In one particularly preferred embodiment, an intein whose protein splicing activity has been blocked by mutation is utilized. The mutant must, however, retain the ability to undergo the N-S shift, thus allowing thioester formation between itself and an N-terminal protein. This thioester can then be nucleophilicly attacked by a thiol reagent or by the N-terminal cysteine of a peptide sequence. For example, by mutating the C-terminal asparagine (asn 198) of an intein from the GyrA gene of *Mycobacterium xenopi* (Telenti, et al., (1997) *J Bacteriol* 179:6378-6382) to an alanine created a

thiol inducible cleavage element. This modified intein cleaved well with thiol reagents that were optimal for the ligation reaction, such as MESNA and thiophenol. Furthermore, optimal thiol reagent and intein combinations can be determined by incubating a precursor protein containing the intein of interest with a wide variety of thiol reagents followed by determination of the extent of cleavage of the precursor protein (Figure 2).

The use of such intein and specific thiol reagents leads to optimal yields and high ligation efficiencies; typically greater than 90% of the N-terminal ligation fragment can be modified.

The ligation methods of the present invention expand the ability to incorporate non-coded amino acids into large protein sequences by generating a synthetic peptide fragment with fluorescent probes, spin labels, affinity tags, radiolabels, or antigenic determinants and ligating this to an *in vivo* expresed protein isolated using a modified intein.

Furthermore, this procedure allows the isolation of cytotoxic proteins by purifying an inactive truncated precursor from a host source, for example bacteria, and generating an active protein or enzyme after the ligation of a synthetic peptide. For example, restriction endonucleases which have not successfully been cloned by traditional

methods may be produced in accordance with the present invention.

Also, the direct ligation procedure allows the ligation of a protein or peptide sequence to another protein or peptide sequence without the use of exogenous thiol reagents. Direct ligation relies on the nucleophilic attack of the N-terminal amino acid of one peptide on the thioester formed between a target protein and an intein (Figure 3).

In summary, a fusion protein can be created using the methods of the present invention that possesses unique properties which, currently, can not be generated genetically.

The Examples presented below are only intended as specific preferred embodiments of the present invention and are not intended to limit the scope of the invention. The present invention encompasses modifications and variations of the methods taught herein which would be obvious to one of ordinary skill in the art.

The references cited above and below are herein incorporated by reference.

EXAMPLE I

Creation of vectors pTXB1 and pTXB2 for ligation:

Asparagine 198 of the *Mxe* GyrA intein (Telenti, et al., (1997) *J Bacteriol.* 179:6378-6382) was mutated to alanine by linker insertion into the *Xmn*I and *Pst*I sites of pmxeMIPTyrXmnSPdel to create pMXP1. The *Xmn*I site was originally introduced into the unmodified *Mxe* GyrA intein sequence by silent mutagenesis. The *Pst*I site was a unique site in the plasmid. The linker was composed of mxe#3 (5'-GGTTCGTCAGCCACGCTACTGGCCTCACCGGTTGATAGCTGCA-3') (SEQ ID NO:1) and mxe#4 (5'-GCTATCAACCGGTGAGGCCAGTAG CGTGGCTGACGAACC-3') (SEQ ID NO:2).

Into pMXP1 another linker composed of mxe#1 (5'-TC GAATCTAGACATATGGCCATGGGTGGCGGCCGCCTCGAGGGCTCTTCC TGCATCACGGGAGATGCA-3') (SEQ ID NO:3) and mxe#2 (5'-CTAG TGCATCTCCCGTGATGCAGGAAGAGCCCTCGAGGCGHGCCGCCACCCA TGGCCATATGTCTAGAT-3') (SEQ ID NO:4) was inserted into the *Xho*I and *Spe*I sites to introduce a multiple cloning site (*Xba*I-*Nde*I-*Nco*I-*Not*I-*Xho*I-*Sap*I) before the *Mxe* GyrA intein (pMXP2).

The 0.6 kilobase *Not*I to *Age*I fragment of pMXP2 was ligated into the same sites in pTYB1 (IMPACT kit, New England Biolabs, Beverly, MA) and the *Nco*I to *Age*I fragment of pMXP2

was cloned into pTYB3 (IMPACT kit, New England Biolabs, Beverly, MA) to create plasmids pTXB1 (see Figure 4 and 5) (SEQ ID NO:5) and pTXB2, respectively. These vectors have a multiple cloning site upstream of the modified *Mxe* GyrA intein-chitin binding domain fusion. This allows the insertion of a target gene of interest inframe with the intein and chitin binding domain (CBD).

Creation of vectors pMYBleu for ligation:

pMYBleu was as described in Chong, et al., (1998), *J. Biol. Chem.* 273:10567-10577. This vector consisted of maltose binding protein upstream of the Sce VMA intein-chitin binding domain. A leucine is present at the -1 position instead of the native residue (which is a glycine).

Purification of Thioester-Tagged Proteins:

Protein purification was as described using the *Sce* VMA intein (Chong, et.al., (1997) *Gene* 192:271-281) with slight modification. ER2566 cells (IMPACT T7 instruction manual from New England Biolabs, Beverly, MA) containing the pTXB vector with the appropriate insert were grown to an OD_{600} of 0.5-0.6 at 37°C at which point they were induced with 0.5 mM IPTG overnight at 15°C. Cells were harvested by centrifugation and lysed by sonication (performed on ice). The

three part fusion protein was bound to chitin beads (10 mL bed volume, Figure 6, lanes 1 and 2) equilibrated in Buffer A (50 mM Tris, pH 7.4, and 500 mM NaCl), and washed with 10 column volumes of Buffer A to remove unbound material.

Cleavage was initiated using a buffer of 50 mM 2-mercaptoethanesulfonic acid (MESNA), 50 mM Tris, pH 8.0 and 100 mM NaCl. Other thiol reagents were also used at other times, such as thiophenol, dithiothreitol, and/or 3-mercaptopropionic acid. After overnight incubation at from 4-25^{o}C protein was eluted from the column (Figure 6 lane 3). This protein contained a thioester at the C-terminus.

Purification of MYB. MYBleu and MXB:

Full length precursor proteins consisting of maltose binding protein-*Sce* VMA intein (N454A)-chitin binding domain (MYB) and maltose binding protein-*Mxe* GyrA (N198A) intein-chitin binding domain (MXB) were purified after induction and sonication, as described above, by applying the sonicated sample to a 10 mL column of amylose resin (New England Biolabs, Beverly, MA). Unbound proteins were washed from the column with 10 column volumes of Buffer A (see purification of thioester-tagged proteins). Bound proteins were eluted with a buffer of 50 mM Tris, pH 8, containing 100 mM NaCl and 10 mM maltose. Fractions were collected and protein

concentrations were determined using the Bio-Rad Protein Assay (Hercules, CA).

Peptide Synthesis:

Peptides for subsequent ligation reactions were synthesized on an ABI model 433A peptide synthesizer utilizing *FastMoc*™ chemistry (Fields, et al., (1991) *Pept Res* 4, 95-101) at a 0.085 mmol scale. Preloaded HMP (p-hydroxymethylphenoxymethyl) polystyrene resins (Applied Biosystems, Foster City, CA) functionalized at 0.5 mmol/g was used in conjunction with Fmoc/NMP chemistry utilizing HBTU amino acid activation (Dourtoglou, et al., (1984) *Synthesis* 572-574; Knorr, et al., (1989) *Tetrahedron Lett* 30, 1927-1930). Fmoc amino acids were purchased from Applied Biosystems (Foster City, CA).

Synthesis proceeded with a single coupling during each cycle. Peptide cleavage from the resin and simultaneous removal of side chain protecting groups was facilitated by the addition of cleavage mixture (Perkin Elmer, Norwalk, CT) consisting of 0.75 g phenol, 0.25 mL 1,2-ethanedithiol, 0.5 mL deionized H_20, and 10 mL TFA. The resin was flushed with nitrogen and gently stirred at room temperature for 3 hours. Following filtration and precipitation into cold (0°C) methyl-t-butyl ether, the precipitate in the ether fraction was

collected by centrifugation. The peptide precipitate was vacuum dried and analyzed by mass spectrometry using a Perceptive Biosystems (Framingham, MA) MALDI-TOF mass spectrometer.

Final purification was by HPLC using a Waters HPLC system with a Lambda-Max Model 481 Multiwavelength detector (set at 214 nm), 500 series pumps and automated gradient controller with a Vydac semi-preparative C18 column. Elution of the peptide was with a 60 minute linear gradient of 6-60% acetonitrile (v/v) in an aqueous solution of 0.1% TFA (v/v).

Protein Cleavage and Ligation Reactions:

Cleavage of MYB and MXB: The precursor protein (1 mg/mL) was incubated overnight at 4°C with or without a thiol reagent (50 mM) in 150 mM Tris, pH 8, containing 100 mM NaCl.

Ligation reactions with MYB and MXB: The precursor protein (1 mg/mL) was treated as described for cleavage except that a 30 amino acid peptide (1 mM final concentration, NH_2-CAYKTTQANKHIIVACEGNPYVPVHFDASV-COOH (SEQ ID NO:6) was also included in the reaction (Figure 2).

Ligation reactions after purification of thioester-tagged proteins: Lyophilized peptides (New England Biolabs, Beverly, MA) were added (to 1 mM final concentration) directly to the thioester-tagged protein freshly isolated from the chitin column. The reaction was allowed to proceed overnight at from 4-25°C. In both ligation procedures the condensation of the reactants is visible on a 10-20% Tricine gel (Figure 6). The ligation reaction was tested in conditions of 5-150 mM Tris or HEPES buffers, 50-1000 mM NaCl, 10 mM Maltose, and pH 6-11 and 0-6 M Urea.

Direct Ligation Reactions:

MYBleu (1 mg/mL) was incubated in 6 M Urea or 1% SDS, pH 7.5-8.5, 50-200 mM NaCl, and 1 mM of a 30 amino acid peptide (NH_2CAYKTTQANKHIVVACEGNPYVPVHFDASV-COOH (SEQ ID NO:6)). The MYBleu was incubated for 0-180 minutes at either 4°C or 100°C prior to the addition of the 30 amino acid peptide. Ligation reactions proceeded overnight at either 4°C or 25°C.

EXAMPLE II

Labeling a target protein: Maltose Binding Protein

Maltose binding protein (MBP, 42 kDa) was isolated as described in Example I above using the IMPACT procedure (IMPACT manual from New England Biolabs, Inc., Beverly, MA) in the presence of MESNA.

A biotinylated peptide possessing an N-terminal cysteine (CDPEK*DS-COOH (SEQ ID NO:9)), in which the biotin was attached to the ε-amino group of the lysine residue) was ligated to the freshly purified target protein as described above. Briefly, 4 μL of biotinylated peptide (10 mM) were mixed with a 36 μL aliquot of the freshly purified MBP sample. The mixture was incubated at 4°C overnight.

Western blots with alkaline phosphatase linked anti-biotin antibody detected the presence of the ligated product but not the unligated target protein (Figure 7). The efficiency of the ligation is typically greater than 90% when MESNA is used for cleavage.

EXAMPLE III

Labeling a target protein: Bst DNA Polymerase I Large Fragment (Bst Pol 1)

Bst DNA Polymerase I large fragment (67 kDa) was isolated as described in Example I above using the IMPACT procedure (IMPACT manual from New England Biolabs, Inc., Beverly, MA) in the presence of MESNA.

A biotinylated peptide possessing an N-terminal cysteine (CDPEK*DS-COOH (SEQ ID NO:9)), in which the biotin was attached to the ε-amino group of the lysine residue) was ligated to the freshly purified target protein as described. Briefly, 4 μL of biotinylated peptide (10 mM) were mixed with a 36 μL aliquot of the freshly purified Bst Pol 1 sample. The mixture was incubated at 4°C overnight.

Western blots with alkaline phosphatase linked anti-biotin antibody detected the presence of the ligated product but not the unligated target protein (Figure 7). The efficiency of the ligation is typically greater than 90% when MESNA is used for cleavage.

EXAMPLE IV

Labeling a target protein: Paramyosin

Paramyosin (29 kDa) was isolated as described in Example I above using the IMPACT procedure (IMPACT manual from New England Biolabs, Inc., Beverly, MA) in the presence of MESNA.

A biotinylated peptide possessing an N-terminal cysteine (CDPEK*DS-COOH (SEQ ID NO:9)), in which the biotin was attached to the ε-amino group of the lysine residue) was ligated to the freshly purified target protein as described. Briefly, 4 μL of biotinylated peptide (10 mM) were mixed with a 36 μL aliquot of the freshly purified paramyosin sample. The mixture was incubated at 4°C overnight.

Western blots with alkaline phosphatase linked anti-biotin antibody detected the presence of the ligated product but not the unligated target protein (Figure 7). The efficiency of the ligation is typically greater than 90% when MESNA is used for cleavage.

EXAMPLE V

Labeling a target protein: *E. coli* Thioredoxin

E. coli thioredoxin (12 kDa) was isolated as described in Example I above using the IMPACT procedure (IMPACT manual from New England Biolabs, Inc., Beverly, MA) in the presence of MESNA.

A biotinylated peptide possessing an N-terminal cysteine (CDPEK*DS-COOH (SEQ ID NO:9)), in which the biotin was attached to the ε-amino group of the lysine residue) was ligated to the freshly purified target protein as described. Briefly, 4 μL of biotinylated peptide (10 mM) were mixed with a 36 μL aliquot of the freshly purified thioredoxin sample. The mixture was incubated at 4°C overnight.

Western blots with alkaline phosphatase linked anti-biotin antibody detected the presence of the ligated product but not the unligated target protein (Figure 7). The efficiency of the ligation is typically greater than 90% when MESNA is used for cleavage.

EXAMPLE VI

Isolation of a cytotoxic protein:

The ligation procedure of Example I was applied to the isolation of a potentially cytotoxic protein. An endonuclease from *Haemophilus parainfluenzae* (*Hpa*I; Ito, et al., (1992) *Nucleic Acids Res* 20:705-709) was generated by ligating an inactive truncated form of the enzyme expressed in *E. coli* (ER2566 cells, New England Biolabs, Inc., Beverly, MA) with the missing amino acids that were synthesized chemically.

The first 223 amino acids of *Hpa*I (full length *Hpa*I is 254 amino acids) were fused in frame with the modified *Mxe* GyrA intein and the CBD. The 223 amino acid *Hpa*I fragment was isolated as described for purification of thioester tagged proteins. The truncated *Hpa*I displayed no detectable enzymatic activity.

A synthetic peptide representing the 31 amino acids needed to complete *Hpa*I was ligated onto the 223 amino acid truncated form of *Hpa*I by the method of Example I.

Enzymatic Assay for *HpaI*:

The activity of the fused *HpaI* was determined by its ability to digest Lambda DNA (New England Biolabs, Beverly, MA). Serial dilutions of ligated or truncated *HpaI*, with the appropriate peptide added to 1 mM, were incubated with 1 μg of Lambda DNA for 1 hour at 37°C in a buffer of 20 mM Tris-acetate, pH 7.9, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM dithiothreitol, and 170 μg/mL BSA (total volume 30 μL). Digestion reactions were visualized on 1% agarose gels permeated with ethidium bromide. One unit of *Hpa* I was defined as the amount of enzyme necessary to digest 1 μg of Lambda DNA in one hour at 37°C.

The newly ligated *HpaI* had a specific activity of 0.5-1.5×10^6 units/mg which correlated well with the expected value of $1\text{-}2 \times 10^6$ units/mg for the full length enzyme.

WHAT IS CLAIMED IS:

1. A method for fusing an expressed protein with a peptide, said method comprising the steps of:
(a) generating at least one C-terminal thioester-tagged target protein;
(b) generating at least one target peptide having a specified N-terminal; and
(c) ligating said target peptide to said target protein.

2. The method of claim 1, wherein said target protein is generated from a first plasmid comprising an intein having N-terminal cleavage activity.

3. The method of claim 2, wherein said intein comprises an intein having a cysteine residue at the N-terminal of the intein.

4. The method of claim 3, wherein said target protein is generated by thiol reagent-induced cleavage of said intein.

5. The method of claim 4, wherein said thiol reagent is selected from the group consisting of MESNA, thiophenol, DTT, ß-mercaptoethanol or derivatives thereof.

6. A fusion protein produced by the method of any one of claims 1-5.

7. A cyclic protein produced by the method of claim 1.

8. A modified intein comprising a mutant Mxe GyrA intein capable of thiol reagent-induced cleavage to produce a thioester at the C-terminal of an adjacent target protein.

9. A method of generating a reactive thioester comprising contacting a thiol reagent selected from the group consisting essentially of MESNA, thiophenol, DTT, ß-mercaptoethanol or derivatives thereof with a precursor comprising a target protein and intein.

10. A method for screening thiol reagents which cleave a target intein comprising the steps of:
   (a) isolating a precursor comprising a protein and a modified intein;
   (b) contacting a thiol reagent with the precursor of step (a);
   (c) determining whether a splicing or cleaving event occurs.

11. The method of claim 10, comprising the further step of determining whether the spliced or cleaved product of step (c) can ligate to a target peptide having an N-temrinal cytokine.

FIG. 1

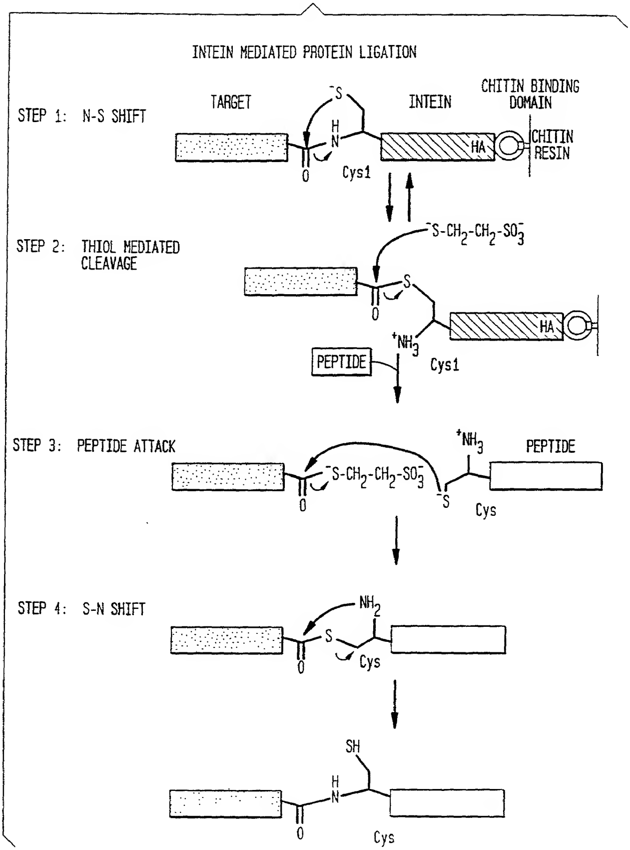

INTEIN MEDIATED PROTEIN LIGATION
STEP 1: N-S SHIFT
TARGET
INTEIN
CHITIN BINDING DOMAIN
HA
CHITIN RESIN
Cys1
STEP 2: THIOL MEDIATED CLEAVAGE
-S-CH2-CH2-SO3-
HA
+NH3
PEPTIDE
Cys1
STEP 3: PEPTIDE ATTACK
+NH3
PEPTIDE
-S-CH2-CH2-SO3-
Cys
STEP 4: S-N SHIFT
NH2
Cys
SH
Cys

FIG. 2

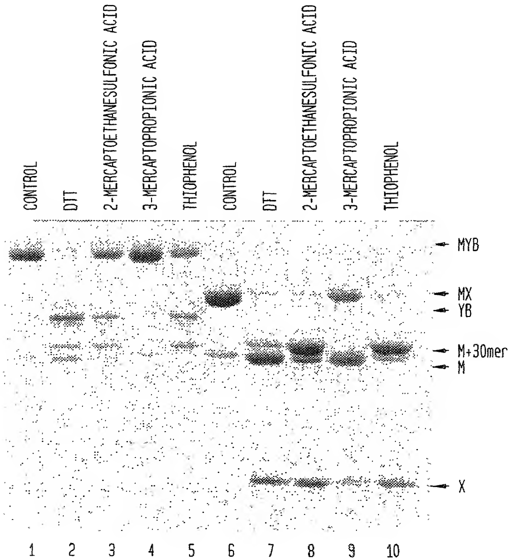

CONTROL
DTT
2-MERCAPTOETHANESULFONIC ACID
3-MERCAPTOPROPIONIC ACID
THIOPHENOL
CONTROL
DTT
2-MERCAPTOETHANESULFONIC ACID
3-MERCAPTOPROPIONIC ACID
THIOPHENOL
MYB
MX
YB
M+30mer
M
X
1
2
3
4
5
6
7
8
9
10

FIG. 3

DIRECT LIGATION REACTON

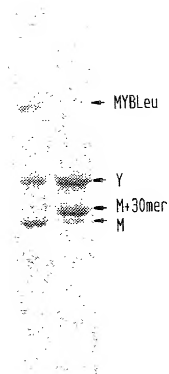
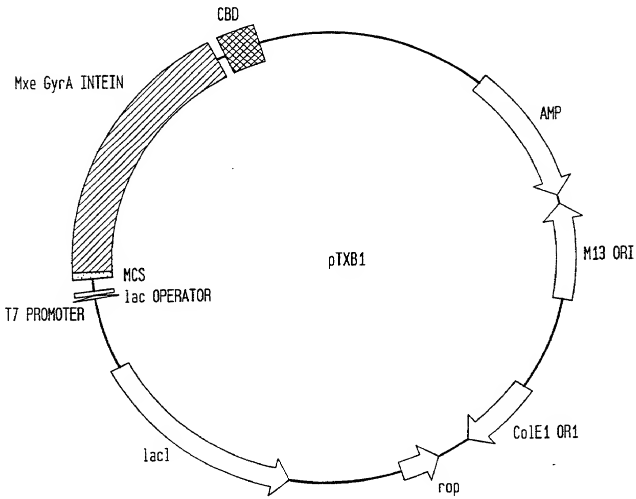


FIG. 4



pTXB1

....Met Ala Met Gly Gly Gly Arg Leu Glu Gly Ser Ser Cysl...Intein

CAT ATG GCC ATG GGT GGC GGC CGC CTC GAG GGC TCT TCC TGC

NdeI NcoI NotI XhoI SapI

FIG. 5A

DNA Sequence of pTXB1 plasmid

140- 997	beta-lactamase (Ap)
1042-1555	M13 origin
2254	ColE1 origin
2626-2814	rop
3376-4455	lacIq
5440-5456	T7 promoter
5440-5459	T7 universal primer (forward)
5457	first nucleotide of the T7 transcript
5459-5483	lac operator
5513-5519	Shine-Dalgarno sequence (T7 gene 10)
5525-5572	Multiple cloning site
5573-6166	Mxe GyrA intein (N198A)
6197-6352	Chitin-binding domain
6375-6497	T7 transcription terminator

TXB1.seq.old Length: 6503 March 17, 1998 11:14 Type: N
Check: 1445 ..

```
  1 AACTACGTCA GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT
 51 GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA
101 CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA
151 ACATTTCCGT GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG
201 TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG
251 TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
301 CCTTGAGAGT TTTCGCCCCG AAGAACGTTC TCCAATGATG AGCACTTTTA
351 AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTGTTGACGC CGGGCAAGAG
401 CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC
451 ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT
501 GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG
551 ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG
601 GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
651 TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG
701 TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA
751 ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT
801 CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG
851 CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC
```

FIG. 5B

901 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC
951 GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA
1001 CTGTCAGACC AAGTTTACTC ATATATACTT TAGATTGATT TACCCCGGTT
1051 GATAATCAGA AAAGCCCCAA AAACAGGAAG ATTGTATAAG CAAATATTTA
1101 AATTGTAAAC GTTAATATTT TGTTAAAATT CGCGTTAAAT TTTTGTTAAA
1151 TCAGCTCATT TTTTAACCAA TAGGCCGAAA TCGGCAAAAT CCCTTATAAA
1201 TCAAAAGAAT AGCCCGAGAT AGGGTTGAGT GTTGTTCCAG TTTGGAACAA
1251 GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAAACCG
1301 TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCCAA ATCAAGTTTT
1351 TTGGGGTCGA GGTGCCGTAA AGCACTAAAT CGGAACCCTA AAGGGAGCCC
1401 CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG
1451 GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG CGCTGGCAAG TGTAGCGGTC
1501 ACGCTGCGCG TAACCACCAC ACCCGCCGCG CTTAATGCGC CGCTACAGGG
1551 CGCGTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA
1601 ATCCCTTAAC GTGAGTTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA
1651 GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT
1701 TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGTTT GCCGGATCAA
1751 GAGCTACCAA CTCTTTTTCC GAAGGTAACT GGCTTCAGCA GAGCGCAGAT
1801 ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA
1851 ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT GTTACCAGTG
1901 GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG
1951 ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA
2001 CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG
2051 CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG
2101 GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC
2151 CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC
2201 TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG
2251 GAAAAACGCC AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC
2301 CTTTTGCTCA CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC

FIG. 5C

2351	CGTATTACCG	CCTTTGAGTG	AGCTGATACC	GCTCGCCGCA	GCCGAACGAC
2401	CGAGCGCAGC	GAGTCAGTGA	GCGAGGAAGC	TATGGTGCAC	TCTCAGTACA
2451	ATCTGCTCTG	ATGCCGCATA	GTTAAGCCAG	TATACACTCC	GCTATCGCTA
2501	CGTGACTGGG	TCATGGCTGC	GCCCCGACAC	CCGCCAACAC	CCGCTGACGC
2551	GCCCTGACGG	GCTTGTCTGC	TCCCGGCATC	CGCTTACAGA	CAAGCTGTGA
2601	CCGTCTCCGG	GAGCTGCATG	TGTCAGAGGT	TTTCACCGTC	ATCACCGAAA
2651	CGCGCGAGGC	AGCTGCGGTA	AAGCTCATCA	GCGTGGTCGT	GCAGCGATTC
2701	ACAGATGTCT	GCCTGTTCAT	CCGCGTCCAG	CTCGTTGAGT	TTCTCCAGAA
2751	GCGTTAATGT	CTGGCTTCTG	ATAAAGCGGG	CCATGTTAAG	GGCGGTTTTT
2801	TCCTGTTTGG	TCACTTGATG	CCTCCGTGTA	AGGGGGAATT	TCTGTTCATG
2851	GGGGTAATGA	TACCGATGAA	ACGAGAGAGG	ATGCTCACGA	TACGGGTTAC
2901	TGATGATGAA	CATGCCCGGT	TACTGGAACG	TTGTGAGGGT	AAACAACTGG
2951	CGGTATGGAT	GCGGCGGGAC	CAGAGAAAAA	TCACTCAGGG	TCAATGCCag
3001	ccgaACGCCA	GCAAGACGTA	GCCCAGCGCG	TCGGCCGCCA	TGCCGGCGAT
3051	AATGGCCTGC	TTCTCGCCGA	AACGTTTGGT	GGCGGGACCA	GTGACGAAGG
3101	CTTGAGCGAG	GGCGTGCAAG	ATTCCGAATA	CCGCAAGCGA	CAGGCCGATC
3151	ATCGTCGCGC	TCCAGCGAAA	GCGGTCCTCG	CCGAAAATGA	CCCAGAGCGC
3201	TGCCGGCACC	TGTCCTACGA	GTTGCATGAT	AAAGAAGACA	GTCATAAGTG
3251	CGGCGACGAT	AGTCATGCCC	CGCGCCCACC	GGAAGGAGCT	GACTGGGTTG
3301	AAGGCTCTCA	AGGGCATCGG	TCGAGATCCC	GGTGCCTAAT	GAGTGAGCTA
3351	ACTTACATTA	ATTGCGTTGC	GCTCACTGCC	CGCTTTCCAG	TCGGGAAACC
3401	TGTCGTGCCA	GCTGCATTAA	TGAATCGGCC	AACGCGCGGG	GAGAGGCGGT
3451	TTGCGTATTG	GGCGCCAGGG	TGGTTTTTCT	TTTCACCAGT	GAGACGGGCA
3501	ACAGCTGATT	GCCCTTCACC	GCCTGGCCCT	GAGAGAGTTG	CAGCAAGCGG
3551	TCCACGCTGG	TTTGCCCCAG	CAGGCGAAAA	TCCTGTTTGA	TGGTGGTTAA
3601	CGGCGGGATA	TAACATGAGC	TGTCTTCGGT	ATCGTCGTAT	CCCACTACCG
3651	AGATATCCGC	ACCAACGCGC	AGCCCGGACT	CGGTAATGGC	GCGCATTGCG
3701	CCCAGCGCCA	TCTGATCGTT	GGCAACCAGC	ATCGCAGTGG	GAACGATGCC
3751	CTCATTCAGC	ATTTGCATGG	TTTGTTGAAA	ACCGGACATG	GCACTCCAGT

FIG. 5D

3801 CGCCTTCCCG TTCCGCTATC GGCTGAATTT GATTGCGAGT GAGATATTTA

3851 TGCCAGCCAG CCAGACGCAG ACGCGCCGAG ACAGAACTTA ATGGGCCCGC

3901 TAACAGCGCG ATTTGCTGGT GACCCAATGC GACCAGATGC TCCACGCCCA

3951 GTCGCGTACC GTCTTCATGG GAGAAAATAA TACTGTTGAT GGGTGTCTGG

4001 TCAGAGACAT CAAGAAATAA CGCCGGAACA TTAGTGCAGG CAGCTTCCAC

4051 AGCAATGGCA TCCTGGTCAT CCAGCGGATA GTTAATGATC AGCCCACTGA

4101 CGCGTTGCGC GAGAAGATTG TGCACCGCCG CTTTACAGGC TTCGACGCCG

4151 CTTCGTTCTA CCATCGACAC CACCACGCTG GCACCCAGTT GATCGGCGCG

4201 AGATTTAATC GCCGCGACAA TTTGCGACGG CGCGTGCAGG GCCAGACTGG

4251 AGGTGGCAAC GCCAATCAGC AACGACTGTT TGCCCGCCAG TTGTTGTGCC

4301 ACGCGGTTGG GAATGTAATT CAGCTCCGCC ATCGCCGCTT CCACTTTTTC

4351 CCGCGTTTTC GCAGAAACGT GGCTGGCCTG GTTCACCACG CGGGAAACGG

4401 TCTGATAAGA GACACCGGCA TACTCTGCGA CATCGTATAA CGTTACTGGT

4451 TTCACATTCA CCACCCTGAA TTGACTCTCT TCCGGGCGCT ATCATGCCAT

4501 ACCGCGAAAG GTTTTGCGCC ATTCGATGGT GTCCCGGATC TCGACGCTCT

4551 CCCTTATGCG ACTCCTGCAT TAGGAAGCAG CCCAGTAGTA GGTTGAGGCC

4601 GTTGAGCACC GCCGCCGCAA GGAATGGTGC ATGCCGCCCT TTCGTCTTCA

4651 AGAATTAATT CCCAATTCCA GGCATCAAAT AAAACGAAAG GCTCAGTCGA

4701 AAGACTGGGC CTTTCGTTTT ATCTGTTGTT TGTCGGTGAA CGCTCTCCTG

4751 AGTAGGACAA ATCCGCCGGG AGCGGATTTG AACGTTGCGA AGCAACGGCC

4801 CGGAGGGTGG CGGGCAGGAC GCCCGCCATA AACTGCCAGG AATTAATTCC

4851 AGGCATCAAA TAAAACGAAA GGCTCAGTCG AAAGACTGGG CCTTTCGTTT

4901 TATCTGTTGT TTGTCGGTGA ACGCTCTCCT GAGTAGGACA AATCCGCCGG

4951 GAGCGGATTT GAACGTTGCG AAGCAACGGC CCGGAGGGTG GCGGGCAGGA

5001 CGCCCGCCAT AAACTGCCAG GAATTAATTC CAGGCATCAA ATAAAACGAA

5051 AGGCTCAGTC GAAAGACTGG GCCTTTCGTT TTATCTGTTG TTTGTCGGTG

5101 AACGCTCTCC TGAGTAGGAC AAATCCGCCG GGAGCGGATT TGAACGTTGC

5151 GAAGCAACGG CCCGGAGGGT GGCGGGCAGG ACGCCCGCCA TAAACTGCCA

5201 GGAATTAATT CCAGGCATCA AATAAAACGA AAGGCTCAGT CGAAAGACTG

09/786009

FIG. 5E

```
5251  GGCCTTTCGT TTTATCTGTT GTTTGTCGGT GAACGCTCTC CTGAGTAGGA
5301  CAAATCCGCC GGGAGCGGAT TTGAACGTTG CGAAGCAACG GCCCGGAGGG
5351  TGGCGGGCAG GACGCCCGCC ATAAACTGCC AGGAATTGGG GATCGGAATT
5401  AATTCCCGGT TTAAACCGGG GATCTCGATC CCGCGAAATT AATACGACTC
5451  ACTATAGGGG AATTGTGAGC GGATAACAAT TCCCCTCTAG AAATAATTTT
5501  GTTTAACTTT AAGAAGGAGA TATAcatatg gctagctcgc gagtcgacgg
5551  cggccgcctc gagggctctt ccTGCATCAC GGGAGATGCA CTAGTTGCCC
5601  TACCCGAGGG CGAGTCGGTA CGCATCGCCG ACATCGTGCC GGGTGCGCGG
5651  CCCAACAGTG ACAACGCCAT CGACCTGAAA GTCCTTGACC GGCATGGCAA
5701  TCCCGTGCTC GCCGACCGGC TGTTCCACTC CGGCGAGCAT CCGGTGTACA
5751  CGGTGCGTAC GGTCGAAGGT CTGCGTGTGA CGGGCACCGC GAACCACCCG
5801  TTGTTGTGTT TGGTCGACGT CGCCGGGGTG CCGACCCTGC TGTGGAAGCT
5851  GATCGACGAA ATCAAGCCGG GCGATTACGC GGTGATTCAA CGCAGCGCAT
5901  TCAGCGTCGA CTGTGCAGGT TTTGCCCGCG GAAAACCCGA ATTTGCGCCC
5951  ACAACCTACA CAGTCGGCGT CCCTGGACTG GTGCGTTTCT TGGAAGCACA
6001  CCACCGAGAC CCGGACGCCC AAGCTATCGC CGACGAGCTG ACCGACGGGC
6051  GGTTCTACTA CGCGAAAGTC GCCAGTGTCA CCGACGCCGG CGTGCAGCCG
6101  GTGTATAGCC TTCGTGTCGA CACGGCAGAC CACGCGTTTA TCACGAACGG
6151  GTTCGTCAGC CACGCTACTG GCCTCACCGG TCTGAACTCA GGCCTCACGA
6201  CAAATCCTGG TGTATCCGCT TGGCAGGTCA ACACAGCTTA TACTGCGGGA
6251  CAATTGGTCA CATATAACGG CAAGACGTAT AAATGTTTGC AGCCCCACAC
6301  CTCCTTGGCA GGATGGGAAC CATCCAACGT TCCTGCCTTG TGGCAGCTTC
6351  AATGActgca ggaaggGGAT CCGGCTGCTA ACAAAGCCCG AAAGGAAGCT
6401  GAGTTGGCTG CTGCCACCGC TGAGCAATAA CTAGCATAAC CCCTTGGGGC
6451  CTCTAAACGG GTCTTGAGGG GTTTTTTGCT GAAAGGAGGA ACTATATCCG
6501  GAT
```

FIG. 6

Hpa I LIGATION

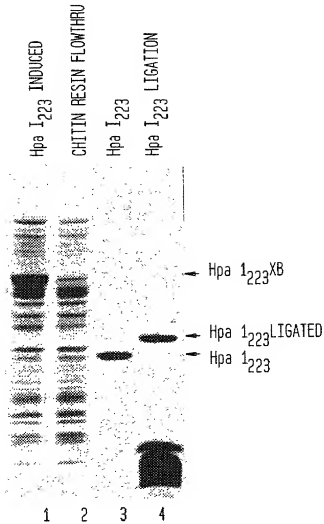
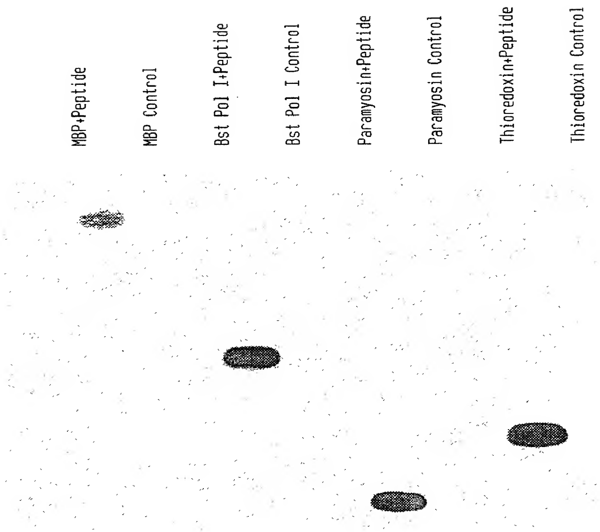


FIG. 7

WESTERN BLOTS OF PROTEINS LIGATED TO A BIOTINYLATED PEPTIDE



New England Biolabs, Inc.
32 Tozer Road
Beverly, MA 01915

DECLARATION
AND POWER OF ATTORNEY
Original Application

Attorney Docket No. NEB-150PUS

As a below named inventor, I hereby declare that:
My residence, post address and citizenship are as stated below next to my name
I believe that I am the original, first and sole inventor (in only one name is listed at 201 below) or an original, first and joint inventor (if plural names are listed at 201-203 below) of the subject matter which is claimed and which a patent is sought on the invention entitled:
INTEIN MEDIATED PEPTIDE LIGATION

which is described and claimed in:
[ ] the attached specification or [ ] the specification in Application Serial No. ________ filed______
(for declaration not accompanying application)
And was amended on________________
if applicable

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendments referred to above. I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).
I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

FOREIGN APPLICATION(S) IF ANY, FILED WITHIN 12 MONTHS PRIOR TO THE FILING DATE OF THIS APPLICATION			
COUNTRY	APPLICATION	DATE OF FILING (day, month, year)	PRIORITY CLAIMED UNDER 35 U.S.C. 119
			YES NO
			YES NO

ALL FOREIGN APPLICATION(S) IF ANY, FILED MORE THAN 12 MONTHS PRIOR TO THE FILING DATE OF THIS APPLICATION			
COUNTRY	APPLICATION	(day, month, year)	PRIORITY CLAIMED UNDER 35 U.S.C. 119
PCT	PCT/US99/22776	30 September 1999	Yes
π			

I hereby claim the benefit under Title 35, United States Code §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

Application Serial No.	Filing Date	Status (Patented, Pending, Abandoned)
60/102,413	30 September 1998	Abandoned

DECLARATION
AND POWER OF ATTORNEY

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorney with full powers of association, substitution and revocation to prosecute this application and transact all business in the Patent and Trademark Office connected therewith:

Gregory D. Williams
(Registration No. 30901)

SEND CORRESPONDENCE TO:

Gregory D. Williams
General Counsel
New England Biolabs, Inc.
32 Tozer Road
Beverly, MA 01915

DIRECT TELEPHONE CALLS TO:

Gregory D. Williams
General Counsel
New England Biolabs, Inc.
Tele: (978) 927-5054; Ext. 292
Fax: (978) 927-1705

2	Full Name of Inventor	Last Name XU	First Name Ming-Qun	Middle Name
0	Residence & Citizenship	City Hamilton	State/Foreign Country Massachusetts MA	Citizenship US
1	Post Office Address	Post Office Address 40 Crescent Road	City/State/Country Hamilton, MA	Zip Code 01984
2	Full Name of Inventor	Last Name EVANS	First Name Thomas	Middle Name C.
0	Residence & Citizenship	City Somerville	State/Foreign Country Massachusetts MA	Citizenship US
2	Post Office Address	Post Office Address 1st Flo 68 Albion Street,	City/State/Country Somerville, MA	Zip Code 02143
2	Full Name of Inventor	Last Name	First Name	Middle Name
0	Residence & Citizenship	City	State/Foreign Country	Citizenship
3	Post Office Address	Post Office Address	City/State/Country	Zip Code
2	Full Name of Inventor	Last Name	First Name	Middle Name
0	Residence & Citizenship	City	State/Foreign Country	Citizenship
4	Post Office Address	Post Office Address	City/State/Country	Zip Code
2	Full Name of Inventor	Last Name	First Name	Middle Name
0	Residence & Citizenship	City	State/Foreign Country	Citizenship
5	Post Office Address	Post Office Address	City/State/Country	Zip Code

206	Full Name of Inventor	Last Name	First Name	Middle Name
	Residence & Citizenship	City	State/Foreign Country	Citizenship
	Post Office Address	Post Office Address	City/State/Country	Zip Code
207	Full Name of Inventor	Last Name	First Name	Middle Name
	Residence & Citizenship	City	State/Foreign Country	Citizenship
	Post Office Address	Post Office Address	City/State/Country	Zip Code
208	Full Name of Inventor	Last Name	First Name	Middle Name
	Residence & Citizenship	City	State/Foreign Country	Citizenship
	Post Office Address	Post Office Address	City/State/Country	Zip Code
209	Full Name of Inventor	Last Name	First Name	Middle Name
	Residence & Citizenship	City	State/Foreign Country	Citizenship
	Post Office Address	Post Office Address	City/State/Country	Zip Code

I hereby further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true and further that these statements were made with the knowledge that willful statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Signature of Inventor 201 Ming-Qun Xu	Date 2/16/01
Signature of Inventor 202 Thomas C. Evans	Date 2/16/01
Signature of Inventor 203	Date
Signature of Inventor 204	Date
Signature of Inventor 205	Date
Signature of Inventor 206	Date
Signature of Inventor 207	Date
Signature of Inventor 208	Date
Signature of Inventor 209	Date

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)__________ of ______________________________ were not present for scanning. (Document title)

☐ Page(s)__________ of ______________________________ were not present for scanning. (Document title)

☐ ***Scanned copy is best available.***

Drawing figures 2, 3, 6, 7 are very dark.